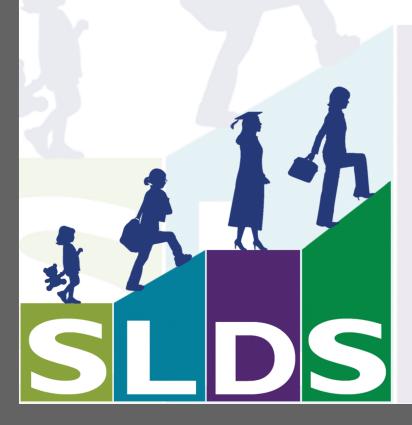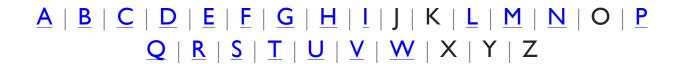# SLDS Glossary

State Support Team

About the Glossary

This glossary was compiled based on input from the Early Childhood Integrated Data System Toolkit Workgroup, the Data Use Standards Workgroup, and the Data Linking Workgroup. It covers terms and definitions commonly used by various workgroups within the Statewide Longitudinal Data System (SLDS) Grant Program, as well as by the Privacy Technical Assistance Center (PTAC). For more information, visit the PTAC Glossary online at http://ptac.ed.gov/glossary.

The glossary is updated regularly. To suggest changes or additions to the glossary, please email support@sst-slds.org.

Within the definitions, terms in **BOLD TEXT** are defined elsewhere in the glossary.

Glossary Contents

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P
Q | R | S | T | U | V | W | X | Y | Z

A

Access

The privilege or assigned permission to use computer data or resources in some manner. It restricts the use and distribution of information, settings, and the general use of a system.

Access Controls

Controls that limit entry to information system resources to authorized **USERS**, programs, processes, or other systems. Components of an access control system include, for example, physical access controls (e.g., locks on doors to a server room), **AUTHENTICATION** systems that verify the identity of a user or client machine attempting to log into a system, and file **ENCRYPTION** that makes data unreadable to anyone who does not possess the cipher key or encryption algorithm.

Accommodation

An alteration of environment, format, or equipment that allows an individual to gain access to content and/or complete assigned tasks.

Accuracy

The closeness of computations or estimates to the exact or true values that the statistics were intended to measure.

Analysis

The process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data.

Anonymization

The process of data de-identification, which produces **DE-IDENTIFIED DATA**, where individual records cannot be linked back to an original student record system or to other individual records from the same source, because they do not include a **RECORD CODE** needed to link the records.

Anonymized

Anonymized data are data that have been **DE-IDENTIFIED** and do not include a **RE-IDENTIFICATION CODE**. In an anonymized data file, the **STUDENT** case numbers in the data records cannot be linked back to the original student record system.

Application

The software designed to perform a specific function directly for the **USER** or, in some cases, for another application program. It may also be called an "application program."

Assessment

The judgment or appraisal of a systematic collection of information about **EDUCATION PROGRAMS,** typically conducted for the purpose of evaluation or improvement.

Assurance Level

The level of confidence in the process used to validate and establish the identity of a person attempting to access an information system. For more guidance on **AUTHENTICATION** assurance levels, see E-authentication Guidance for Federal Agencies (OMB M04-04) and Electronic Authentication Guideline (NIST SP 800-63-1).

Attendance

Defined by **FERPA** regulations to include, but not be limited to: (a) attendance in person or by paper correspondence, videoconference, satellite, Internet, or other electronic information and telecommunications technologies for **STUDENTS** who are not physically present in the classroom; and (b) the period during which a person is working under a work-study program. See also **DATES OF ATTENDANCE**.

Attribute Disclosure

Disclosure occurring when data in a student-level file or aggregate data in tabulations reveal sensitive information about a **STUDENT**.

Audit

A personal or computerized review process that accounts for the adequacy, effectiveness, security, and overall functionality of a data activity. Note: This use of "audit" varies from the audit and evaluation exception under **FERPA**.

Authentication (Single and Multifactor)

A mechanism that an electronic system uses to identify and validate the identity of **USERS** with the required degree of confidence that the user is who he or she purports to be. Authentication is accomplished through the use of one or more "factors," which correspond to things that the user knows (like a password), something that they possess (like a **SECURITY TOKEN**), or something they are (like a fingerprint). Authentication should not be confused with **AUTHORIZATION**, which is the process of granting individuals access to system resources based on their identity. More guidance on identity credentials is available from the National Institute for Standards of Technology publication NIST SP 800-103.

Authenticator

A "secret that creates the binding between credentials and its presenter," as described in the National Standards of Technology publication NIST SP 800-103. Authenticators can take the form of information, such as passwords or PINs; hardware, such as key fobs or smart cards; or some digital form, such as digital signatures and certificates.

Authorization

The permission to **ACCESS** non-public information or use equipment that is either fully or partially restricted, and the process of establishing what actions the authorized **USER** is permitted to perform.

Authorized Disclosure

See **UNAUTHORIZED DISCLOSURE**, **DISCLOSURE**, and **IDENTITY DISCLOSURE**.

Authorized Representative

Any entity or individual designated by a state or local educational authority or an agency headed by an official listed in 34 CFR §99.31(a)(3) [i.e., Comptroller General of U.S., U.S. Attorney General, U.S. Secretary of Education, and state or local educational authorities] to conduct—with respect to federal- or state-supported **EDUCATION PROGRAMS**—any **AUDIT** or evaluation, or any compliance or enforcement activity in connection with federal legal requirements that relate to these programs.

B

Baseline

The level of performance at the start of **DATA COLLECTION** or program that can be used to measure change in indicators in the future.

Bias

The values or beliefs of individuals conducting research and evaluation that can influence the results, including the information to be collected, how it is collected, which **ANALYSES** are optimal, and how the results should be presented. Bias needs to be recognized, controlled for, and mitigated to minimize potential errors and/or subjectivity in results that can occur by selecting a value, outcome, or answer over others based on values or beliefs—whether intentional or unintentional.

Biometric Record

Defined by **FERPA** regulations as one or more measurable biological or behavioral characteristics that can be used for automated recognition of an individual. Examples include fingerprints, retina and iris patterns, voiceprints, DNA sequence, facial characteristics, and handwriting.

Blurring

A **DISCLOSURE LIMITATION METHOD** that is used to reduce the precision of the **DISCLOSED** data to minimize the certainty of individual **IDENTIFICATION**. There are many possible ways to implement blurring, such as by converting continuous **DATA ELEMENTS** into categorical data elements (e.g., creating categories that subsume unique cases), aggregating data across small groups of respondents, and reporting rounded values and ranges instead of exact counts to reduce the certainty of identification. Another approach involves replacing an individual's actual reported value with the average group value; it may be performed on more than one variable with different groupings for each variable.

Brute-Force Attack

A type of malicious attack against a system in which the attacker repeatedly attempts to gain access by presenting all possible combinations of **ACCESS** credentials until a match is found. A hacker attempting to gain access to a system by guessing all possible combinations of characters in a password is an example of a brute-force attack.

Business Intelligence

An umbrella term that includes the **APPLICATIONS**, infrastructure and tools, and best practices that enable **ACCESS** to and analysis of information to improve and optimize decisions and performance.

Business Logic

Custom rules or algorithms that handle the exchange of information between a **DATABASE** and **USER INTERFACE**. Business logic essentially consists of **BUSINESS RULES**, which are policies that govern various aspects of a business, and **WORKFLOWS**, which are sequences of steps that specify in detail the flow of information or data. Also known as business rules or domain logic. An example of business logic is how average **ATTENDANCE** is calculated from attendance sheets.

Business Requirements

Constraints, demands, necessities, needs, or parameters—defined by the business community or customers—that must be met or satisfied, usually within a certain timeframe.

Business Rules

Statements that impose constraints on the selection, relationships, and structure of the **DATA ELEMENTS** in a **DATABASE**.

C

Causal

A claim used to purport a cause-and-effect relationship based on data analysis. Causal claims can only be made when results are statistically significant within experimental design and randomized control trials. Typically, analytic results render **CORRELATIONS** rather than causal claims.

Centralized System

In a centralized data system, all participating source systems copy their data to a single, centrally located data repository where they are organized, integrated, and stored using a common **DATA STANDARD**. Data in a **P-20W+** centralized **SLDS** are periodically matched, integrated, and loaded into a central repository. **USERS** query the system and can **ACCESS** the data that they have been **AUTHORIZED** to view and use.

Coarsening

A **DISCLOSURE** limitation technique to preserve an individual's data by reducing the level of detail used to report some variables. Examples of this technique include recoding continuous variables into intervals, recoding categorical data into broader intervals, and top- or bottom-coding the ends of continuous distributions.

Code (or Option) Set

The list of codes in an option set defining the limited set of value options allowed for a **DATA ELEMENT**. The code set contains numeric or alphanumeric "codes" without spaces or special characters for machine readability. **COMMON EDUCATION DATA STANDARDS (CEDS)** option sets provide a list of codes (code set) and corresponding human-readable descriptions/definitions. "Code set" and "option set" are used interchangeably by some education agency IT professionals.

Common Education Data Standards (CEDS)

A national collaborative effort to develop voluntary, common **DATA STANDARDS** for a key set of education **DATA ELEMENTS** to streamline the exchange, comparison, and understanding of data within and across **P-20W+** institutions and sectors.

Confidentiality

The responsibility of a person who has **ACCESS** to another individual's personal data not to share the data without consent.

Configuration Management Policy

The management of security features through control of changes made to hardware, software, firmware, and security documentation throughout the lifecycle of an information system. Also referred to as secure configuration management policy.

Congruent

Matching or in agreement with something.

Correlation

A claim that shows the degree of relationship that exists between two things.

Critical Reflection

Disciplined thinking that is informed by evidence.

Cross Sectional

Data that are collected from multiple sectors or sources at the same point in time and then brought together for **ANALYSIS**. Cross-sectional data are useful for a snapshot at one point in time typically to describe **TRENDS** across all groups and identify differences among subgroups.

Cross-Site Request Forgery

A type of malicious exploit where an attacker gains **ACCESS** to and executes unauthorized commands on a target web **APPLICATION** (e.g., a web interface for a network device or a web email client) via the browser of an already authenticated **USER**. The attack is accomplished by tricking a validated user who has logged in and has a session cookie stored in the browser into opening an email message or visiting a webpage with embedded malicious content.

Cross-Site Scripting (XSS)

A type of computer security vulnerability that uses malicious script embedded in an otherwise benign and trusted web **APPLICATION** to gather **USER** data. When the script is executed (e.g., when a user clicks on a compromised link in an email message or reads an infected forum post), sensitive user data can be accessed by the attacker.

Cryptographic Hash Algorithm

A well-defined computational procedure that takes variable inputs (e.g., an individual's name or a password in plain text) and produces a fixed-size hash value (also known as message digest), such that any accidental or intentional change of the input (e.g., a **USER** mistyping a password) would produce a different hash value.

D

Dashboard

A visual representation of data that helps **USERS** identify correlations, **TRENDS**, outliers (anomalies), patterns, and business conditions. A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance.

Data Acquisition

The process of collecting or gaining **ACCESS** to and organizing information.

Data Aggregation

A process in which information is gathered and expressed in a summary form, for purposes such as statistical analysis to see group-level performance or information based on specific **VARIABLES** and/or to protect individuals' identity.

Data Audit

The reviewing of data to assess its quality or utility for a specific purpose. Auditing data, unlike auditing finances, involves looking at key metrics other than quantity to create conclusions about the properties of a dataset.

Data Breach

The intentional or unintentional release of secure information to an untrusted environment.

Data Cleansing

The process of amending or removing data in a **DATABASE** that is incorrect, incomplete, improperly formatted, or duplicated. Also known as "data scrubbing."

Data Collection

The process of gathering information in a systematic fashion for analysis and reporting.

Data Definition

Language for describing data or information structures.

Data Dictionary

An agreed-upon set of clearly and consistently defined **DATA ELEMENTS**, **DATA DEFINITIONS**, and attributes. A data dictionary helps an organization maintain consistency in its information systems. **DATABASE** users and managers refer to a data dictionary to find out where specific data are located, whether they were reported correctly, how to use them appropriately, and what their values mean.

Data Documentation

Information on the context of **DATA COLLECTION**, including collection methods (e.g., sampling, instruments, technology used); data sources; **DATA VALIDATION** and modification; and data **CONFIDENTIALITY**, **ACCESS**, and use conditions.

Data Element

An atomic unit of data that has precise meaning or precise semantics that can be defined and/or measured.

Data Governance

The overall management of the availability, usability, integrity, quality, and security of data. Data governance is both an organizational process and a structure. It establishes responsibility for data, organizing program area/agency staff to collaboratively and continuously improve **DATA QUALITY** through the systematic creation and enforcement of policies, roles, responsibilities, and procedures.

Data Integration

The combination of technical and business processes used to combine data from disparate sources into meaningful and valuable information. A complete data integration solution encompasses discovery, cleansing, monitoring, transforming, and delivery of data from a variety of sources.

Data Linkage

A merging of information from two or more sources of data to consolidate facts about an individual or an event that are not available in any separate, single record.

Data Matching

A process of comparing information in two or more datasets in order to determine the likelihood of **DATA ELEMENTS** in each dataset representing the same entity.

Data Model

A conceptual, logical, or physical representation of the **DATA ELEMENTS**, entities, and relationships.

Data Products

Any representation that depicts or summarizes data and analyses (e.g., charts, graphs, tables, documents, reports, briefs, data **DASHBOARDS**).

Data Quality

A multi-dimensional measurement of the adequacy of a particular datum or datasets based on a number of dimensions including, but not limited to, accuracy, completeness, consistency, and timeliness.

Data Security

The means of ensuring that data are kept safe from corruption and that **ACCESS** to the data is suitably controlled. The primary goal of any information and technology security system is to protect information and system equipment without unnecessarily limiting access to authorized **USERS** and functions.

Data Sharing Agreement

A formal contract that clearly documents what data are being shared and how the data can be used.

Data Standard

A documented agreement on representations, formats, and definitions of common data. Data standards are intended to improve the quality and share-ability of education data.

Data Stewards

Managers and administrators within an organization who are responsible for implementing **DATA GOVERNANCE** policies and standards and maintaining **DATA QUALITY** and **DATA SECURITY**.

Data Stewardship

A comprehensive approach to data management to ensure quality, integrity, accessibility, and security of the data.

Data Store

A repository for storing, managing, and distributing datasets on an enterprise level. It is a broad term that incorporates all types of data that are produced, stored, and used by an organization. The term references data that are at rest and used by one or more data-driven **APPLICATIONS**, services, or individuals.

Data System Architecture

Models, policies, rules, or standards that govern which data are collected and how they are stored, arranged, and put to use in a **DATABASE** system and/or organization.

Data System Enhancement

A noteworthy improvement included in a new version of a product. The term is also sometimes used to distinguish an improvement (enhancement) of some existing product capability from a totally new capability.

Data Use

The process by which people examine and make sense of data to inform decisions and actions.

Data Validation

The inspection of data for completeness and reasonableness, and the correction of errors (e.g., incorrect values).

Data Visualization

A general term that describes any effort to help people understand the significance of data by placing them in a visual context. Data visualization tools include standard charts and graphs and displaying data in more sophisticated ways such as infographics; dials and gauges; geographic maps; sparklines; heat maps; and detailed bar, pie, and fever charts.

Data Warehouse

A subject-oriented, integrated, time-variant, and non-volatile collection of data in support of management's decisionmaking process.

Database

A data structure that stores organized information.

Dates of Attendance

Defined by **FERPA** regulations as the period of time during which a **STUDENT** attends or attended an **EDUCATIONAL AGENCY OR INSTITUTION**. Examples of dates of attendance include an academic year, a spring semester, or a first quarter. The term does not include specific daily records of a student's attendance at an educational agency or institution. See also **ATTENDANCE**.

De-Identification Strategy

See **DISCLOSURE LIMITATION METHOD**.

De-Identified Data

Records that have a **RE-IDENTIFICATION CODE** and have enough **PERSONALLY IDENTIFIABLE INFORMATION** removed or obscured so that the remaining information does not identify an individual and there is no reasonable basis to believe that the information can be used to identify an individual. The re-identification code may allow the recipient to match information received from the same source.

Descriptive Statistics

A type of **ANALYSIS** used to describe, show, or summarize data meaningfully so that patterns or **TRENDS** can emerge. Results can include frequencies, mean, sums, etc.

Direct Identifier

Information that relates specifically to an individual, such as the individual's residence, name, address, Social Security number or other identifying number or code, telephone number, email address, or **BIOMETRIC RECORD**. See also **INDIRECT IDENTIFIER**.

Directory Information

Information contained in the **EDUCATION RECORDS** of a **STUDENT** that would not generally be considered harmful or an invasion of privacy if disclosed. Typically, directory information includes information such as name, address, telephone listing, date and place of birth, participation in officially recognized activities and sports, and **DATES OF ATTENDANCE**. A school may disclose directory information to third parties without consent if it has given public notice of the types of information which it has designated as directory information, the parent's or **ELIGIBLE STUDENT'S** right to restrict the **DISCLOSURE** of such information, and the period of time within which a parent or eligible student has to notify the school in writing that he or she does not want any or all of those types of information designated as directory information ([34 CFR § 99.3 and 34 CFR § 99.37](#)).

Disciplinary Action or Proceeding

Defined by **FERPA** regulations as the investigation, adjudication, or imposition of sanctions by an **EDUCATIONAL AGENCY OR INSTITUTION** with respect to an infraction or violation of the internal rules of conduct applicable to **STUDENTS** of the agency or institution.

Disclosure

ACCESS to or the release, transfer, or other communication of **PERSONALLY IDENTIFIABLE INFORMATION** by any means ([34 CFR §99.3](#)). Disclosure can be **AUTHORIZED**, such as when a parent or an **ELIGIBLE STUDENT** gives written consent to share **EDUCATION RECORDS** with an authorized **PARTY** (e.g., a researcher). Disclosure can also be **UNAUTHORIZED** or inadvertent (accidental). An unauthorized disclosure can happen due to a **DATA BREACH** or a loss, and an accidental disclosure can occur when data released in public aggregate reports are unintentionally presented in a manner that allows individual **STUDENTS** to be identified.

Disclosure Avoidance

The efforts made to **DE-IDENTIFY** the data in order to reduce the risk of **DISCLOSURE** of **PERSONALLY IDENTIFIABLE INFORMATION**. A choice of the appropriate **DE-IDENTIFICATION STRATEGY** (also referred to as **DISCLOSURE LIMITATION METHOD**) depends on the nature of the data release, the level of protection offered by a specific method, and the usefulness of the resulting data product. The two major types of data release are aggregated data (such as tables showing numbers of enrolled **STUDENTS** by race, age, and sex) and microdata

(such as individual-level student assessment results by grade and school). Several acceptable de-identification methods exist for each type of data (see **DISCLOSURE LIMITATION METHOD** for more details).

Disclosure Limitation Method

A general term referring to a statistical technique used to manipulate the data prior to release to minimize the risk of inadvertent or **UNAUTHORIZED DISCLOSURE** of **PERSONALLY IDENTIFIABLE INFORMATION**. Also known as a disclosure avoidance method.

E

Early Childhood Education Program

An early childhood education program is defined as

    (a) a **HEAD START** program or an **EARLY HEAD START** program carried out under the Head Start Act, including a migrant or seasonal Head Start program, an Indian Head Start program, or a Head Start program or an Early Head Start program that also receives state funding;

    (b) a state-licensed or -regulated child care program; or

    (c) a program that

        (1) serves children from birth through age 6 that addresses the children's cognitive (including language, early literacy, and early mathematics), social, emotional, and physical development; and

        (2) is a state prekindergarten program; a program authorized under **PART B 619** or **PART C** of the Individuals with Disabilities Education Act; or a program operated by a local **EDUCATIONAL AGENCY**.

Early Childhood Integrated Data System (ECIDS)

A data system that collects, integrates, maintains, stores, and reports information from early childhood programs across multiple agencies within a state that serve children and families from birth to age 8. Typically, the data included in an ECIDS are related to the individual child, the child's family, the classroom, the program/providers, and other services that provide comprehensive care and education for young children.

Early Head Start

A federal program that serves pregnant women, infants, and toddlers. Early Head Start programs are available to the family until the child turns 3 years old and is ready to transition into **HEAD START** or another prekindergarten program. Early Head Start helps families care for their infants and toddlers through early, continuous, intensive, and comprehensive services.

Education Program

Any program principally engaged in the provision of education, including, but not limited to, **EARLY CHILDHOOD EDUCATION**, elementary and secondary education, postsecondary education, special education, job training, career and technical education, and adult education, and any program that is administered by an **EDUCATIONAL AGENCY OR INSTITUTION**.

Education Records

Records that are directly related to a **STUDENT** and are maintained by an **EDUCATIONAL AGENCY OR INSTITUTION** or by a **PARTY** acting for the agency or institution.

Educational Agency or Institution

Any public or private agency or institution to which funds have been made available under any program administered by the U.S. Secretary of Education, if the educational institution provides educational services or instruction, or both, to **STUDENTS**; or the educational agency is authorized to direct and control public elementary or secondary institutions, or **INSTITUTIONS OF POSTSECONDARY EDUCATION**.

Eligible Student

Defined by **FERPA** as a **STUDENT** who has reached 18 years of age or is attending an **INSTITUTION OF POSTSECONDARY EDUCATION** at any age. This means that, at the secondary level, once a student turns 18, all the rights that once belonged to his or her parents transfer to the student. However, a secondary school or postsecondary institution may still provide an eligible student's parents with access to **EDUCATION RECORDS** without the student's consent if the student is claimed as a dependent for IRS tax purposes.

Encryption

The process of transforming information using a **CRYPTOGRAPHIC** algorithm (called a cipher) to make it unreadable to anyone except those possessing special knowledge, usually referred to as an encryption/decryption key.

Enterprise Architecture

A conceptual tool that assists organizations with understanding their own structure and the way they work. Enterprise architecture provides a map of the enterprise and is a route planner for business and technology change.

F

Family Educational Rights and Privacy Act (FERPA)

A federal law (20 U.S.C. § 1232g; 34 CFR Part 99) that protects the privacy of student **EDUCATION RECORDS**. The law applies to all schools that receive funds under an applicable program of the U.S. Department of Education. FERPA gives parents certain rights with respect to their children's education records, including the right to have **ACCESS** to their children's education records, the right to seek to have the records amended, and the right to consent to the **DISCLOSURE** of **PERSONALLY IDENTIFIABLE INFORMATION** from education records, except as provided by law. These rights transfer to the student when he or she reaches the age of 18 or attends a school beyond the high school level. Students to whom the rights have transferred are **ELIGIBLE STUDENTS**.

Federated System

In a federated data system, individual source systems maintain control over their own data but agree to share some or all of this information with other participating systems upon request. System **USERS** submit queries via a shared intermediary interface that then searches the independent source systems. Data are queried from source systems, and records are matched to fulfill a data requestor's information needs. The linked data are not stored by the system, but rather are removed once cached and delivered. The individual sources of data store and secure their data, and they provide them to the system only upon request.

Financial Aid

Defined by **FERPA** regulations as a payment of funds provided to an individual (or a payment in kind of tangible or intangible property to the individual) that is conditioned on the individual's **ATTENDANCE** at an **EDUCATIONAL AGENCY OR INSTITUTION**.

Findings

The principal outcomes of a research project examining data, and what the project data suggested, revealed, or indicated.

Formative Assessment or Evaluation

Describes a method whereby the information collected is used to improve or inform progress before or during a learning opportunity or program implementation in order to improve performance, outcome, or design.

G

Generalizability

The ability to infer from specific facts, statistics, examples, or the like to a broader group.

Growth Measures

Measures comparing the relative change in performance of one person on a specific test or outcome measure to themselves or others using previous data (e.g., prior test scores). Indicators of growth can usually be aggregated for groups of people with common characteristics (e.g., race/ethnicity, subject, grade, socioeconomic status, gender). Some examples of growth measures include student growth percentiles (SGPs), value-added modeling, and transition table outcomes.

H

Harm

Intentional or unintentional physical or mental damage or injury to subjects as a result of using subjects' data.

Head Start

A federal program that promotes the school readiness of children ages birth to age 5 from low-income families by enhancing their cognitive, social, and emotional development.

Health Insurance Portability and Accountability Act (HIPAA)

The 1996 law that sets requirements for electronic health transactions (billing and payment) and establishes the basis for the HIPAA Privacy Rule. The HIPAA Privacy Rule standards implemented by the U.S. Department of Health and Human Services address the use and **DISCLOSURE** of individuals' health information—called "protected health information" by organizations subject to the Privacy Rule ("covered entities")—as well as standards for individuals' privacy rights to understand and control how their health information is used. HIPAA requirements are set through a suite of regulations, collectively known as HIPAA Administrative Simplification. They include rules on Privacy, Enforcement, Security and Transactions, and Code Sets.

High-Quality Data

Data that are generally agreed upon as trustworthy and are gathered through a reliable instrument or process that is **VALID FOR USE**.

Human Subjects

Defined by U.S. Department of Health and Human Services regulations under 45CFR46.102(f) as a living individual about whom an investigator conducting research obtains (1) data through intervention or interaction with the individual; or (2) identifiable private information.

I

Identifiable Form

Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either **DIRECT IDENTIFIERS** or **INDIRECT IDENTIFIERS**.

Identification

The process that a system uses to recognize a valid **USER'S** asserted identity. It is the first step in the **AUTHENTICATION** and **AUTHORIZATION** process where the user requesting **ACCESS** is asserting that he or she is a valid user. This differs from the authentication process during which the user provides a proof, or factors that prove, that the user really is the person he or she claims to be.

Identity Disclosure

Occurs when data in a student-level file or aggregate data in tabulations allow the data **USER** to identify a **STUDENT**.

Indirect Identifier

Information that can be combined with other information to identify specific individuals, including, for example, a combination of gender, birth date, geographic indicator, and other descriptors. Other examples of indirect identifiers include place of birth, race, religion, weight, activities, employment information, medical information, education information, and financial information. See also **DIRECT IDENTIFIER**.

Inferential Statistics

A type of **ANALYSIS** where a sample of data is used to generalize or infer something about a larger population. This analysis usually requires a robust sampling design and the minimization of sampling **BIAS** to reduce error and to produce objective, **VALID-FOR-USE** results.

Information Security

The protection of information and information systems from unauthorized **ACCESS**, use, **DISCLOSURE**, disruption, modification, or destruction in order to provide **CONFIDENTIALITY**, integrity, and availability.

Institution of Postsecondary Education

Defined by **FERPA** regulations as an institution that provides education to **STUDENTS** beyond the secondary school level.

Instructional Material

Defined by the Protection of Pupil Rights Amendment (PPRA) as all material provided to a **STUDENT**, regardless of format, including printed or representational materials, audio-visual materials, and materials in electronic or digital formats (such as materials accessible through the Internet). The term does not include academic tests or academic assessments.

Interoperability

A set of rules, definitions, and transport processes that enable different software systems to share information and work together.

L

Law Enforcement Unit

Defined by **FERPA** regulations as any individual, office, department, division, or other component of an **EDUCATIONAL AGENCY OR INSTITUTION**, such as a unit of commissioned police officers or non-commissioned security guards, that is officially authorized or designated by that agency or institution to (1) enforce any local, state, or federal law, or refer to appropriate authorities a matter for enforcement of any local, state, or federal law against any individual or organization other than the agency or institution itself; or (2) maintain the physical security and safety of the agency or institution. A component of an educational agency or institution does not lose its status as a law enforcement unit if it also performs other, non-law enforcement functions for the agency or institution, including investigation of incidents or conduct that constitutes or leads to a **DISCIPLINARY ACTION** or proceedings against the **STUDENT**.

Longitudinal Data

Data that are collected on a specific group of people over a period of time. These data are useful for understanding change over time.

M

Masking

A **DISCLOSURE LIMITATION METHOD** that is used to "mask" the original values in a dataset to achieve data privacy protection. This general approach uses various techniques, such as data **PERTURBATION**, to replace sensitive information with realistic but inauthentic data or modifies original data values based on pre-determined masking rules (e.g., by applying a transformation algorithm). The purpose of this technique is to retain the structure and functional usability of the data, while concealing information that could lead to the **IDENTIFICATION**, either directly or indirectly, of an individual **STUDENT**. Masked data are used to protect individual privacy in public reports and can serve as a useful alternative for occasions when the real data are not required, such as **USER** training or software demonstration. Specific masking rules may vary depending on the sensitivity level of the data and organizational data **DISCLOSURE** policies.

Measure

A systematic unit or standard based on the assignment of symbols, letters, or numbers to properties of **VARIABLES** according to rules. There are four levels of measures: nominal,[1] ordinal,[2] interval,[3] and ratio.[4]

Memorandum of Understanding (MOU)

A formal agreement between two or more parties. Note: Many times an MOU is required under **FERPA** when linking data.

Multiple Measures

A variety of measures—such as the results from standardized tests, classroom **ASSESSMENTS**, tasks and projects, grades, and teacher evaluations—that could be used to provide a complete picture of a **STUDENT'S** academic achievement, a school's or district's performance, or a group's performance.

---

[1] Nominal: A level of measure where assigned values have no true meaning without the use of a key or crosswalk. For example, to indicate lunch status 0= not free or reduced, 1 = reduced status, 2 = free status.

[2] Ordinal: A level of measure where the assigned value shows a rank or order, but the distance between the values has no meaning. For example, the student ranked as 1 received the highest score of 97; the student ranked as 2 received a 95; and the student ranked as 3 received an 89.

[3] Interval: A level of measure where the assigned value shows an order and the distance between the values does have meaning. For example, the temperature differs by 10 whether between 20 and 30 degrees or between 75 and 85 degrees, and the 10 is interpretable in both cases.

[4] Ratio: A level of measure where there is a scale with an absolute value, which allows for the use of fractions and ratios. For example, if 25 students were proficient last year and 50 are proficient this year, twice as many students are proficient this year as compared to last year.

N

Network Security Mechanisms

The security products and policies used by network security personnel to prevent and monitor unauthorized **ACCESS** misuse, modification, or denial of the information system and network resources. For example, anti-virus and email security software are network security mechanisms.

P

P-20W+

Refers to data from prekindergarten (early childhood), K12, and postsecondary through post-graduate education, along with workforce and other outcomes data (e.g., public assistance and corrections data). The specific agencies and other organizations that participate in the P-20W+ initiative vary from state to state.

Part B 619

A program that serves children ages 3 to 5 by providing special services to children diagnosed with developmental delays and disabilities who are eligible under the Individuals with Disabilities Education Act.

Part C

A $436 million program administered by states that serves infants and toddlers through age 2 with developmental delays or who have diagnosed physical or mental conditions with high probabilities of resulting in developmental delays.

Party

Defined by **FERPA** regulations as an individual, agency, institution, or organization.

Perimeter Security Mechanisms

The specific security policies and products used at the network perimeter, which is defined as the boundary between the private, locally managed and operated side of the network and the public side of the network. For example, a firewall and an intrusion detection system are perimeter security mechanisms.

Personal Information Collected from Students

A Protection of Pupil Rights Amendment (PPRA) term referring to individually identifiable information including a **STUDENT'S** or parent's first and last name; a home or other physical address (including street name and the name of the city or town); a telephone number; or a Social Security number collected from any elementary or secondary school student.

Personally Identifiable Information (PII)

Defined in **FERPA** as including, but not limited to, a **STUDENT** name; the name of the student's parents or other family members; the address of the student or student's family; a personal identifier such as the student's date of birth, place of birth, and mother's maiden name; other information that, alone or in combination, is linked or linkable to a specific student that would allow a reasonable person in the school

community, who does not have personal knowledge of the relevant circumstances, to identify the student with reasonable certainty; and information requested by a person who the **EDUCATIONAL AGENCY OR INSTITUTION** reasonably believes knows the identity of the student to whom the **EDUCATION RECORD** relates.

Personally Identifiable Information for Education Records

A **FERPA** term referring to identifiable information that is maintained in **EDUCATION RECORDS** and includes **DIRECT IDENTIFIERS**, such as a **STUDENT'S** name or identification number, **INDIRECT IDENTIFIERS**, such as a student's date of birth, or other information that can be used to distinguish or trace an individual's identity either directly or indirectly through linkages with other information.

Perturbation

A **DISCLOSURE LIMITATION METHOD** that involves making small changes to the data to prevent **IDENTIFICATION** of individuals from unique or rare population groups. Perturbation is a data **MASKING** technique in that it is used to "mask" the original values in a data set to avoid **DISCLOSURE**.

Planning and Management

Planning is deciding in advance what is to be done, when, where, how, and by whom to achieve the **PURPOSE AND VISION** for the data system. It includes establishing strategies, objectives, policies, and procedures. Management is the oversight of the execution of a plan, including necessary adjustments over time to reflect changes in context, needs, and resources.

Portability

In relation to software, a measure of how easily an **APPLICATION** can be transferred from one computer environment to another with few or no modifications.

Predictive

A type of **DATA ANALYSIS** that uses current and historical facts to make predictions about the future.

Privacy

An individual's right to have his or her personal information kept private. This right is balanced by the need for **DATA COLLECTION** and dissemination.

Psychologically Safe

A shared belief that the group or context is safe for interpersonal risk-taking involved in the collection, analysis, and sharing of data.

Purpose and Vision

The purpose and vision statements communicate the data system's reason for being, what it aims to produce, and how it contributes to the long-term policy and program goals of the state.

Q

Qualitative Data

Data that are not expressed numerically.

Quantitative Data

Data that are expressed numerically.

R

Re-Identification Code

A re-identification code enables an authorized researcher to return to the source of **DE-IDENTIFIED DATA** and match the de-identified data to the source records.

Record Code

A unique descriptor that can be used to match individual-level records across **DE-IDENTIFIED DATA** files from the same source (e.g., for the purposes of comparing performance of individual **STUDENTS** over time).

Redaction

The process of expunging (striking out or deleting) sensitive data from education and early intervention records prior to **DISCLOSURE** in a way that meets established disclosure requirements applicable to the specific data disclosure occurrence (e.g., removing or obscuring **PERSONALLY IDENTIFIABLE INFORMATION** from published reports to meet federal, state, and local privacy laws, as well as organizational data disclosure policies).

Reliability of Data

The degree to which the results are replicable, accurate, and complete given the specifications of a particular collection.

Reliable Data

Data that are collected in a consistent manner with standard **BUSINESS RULES** in place allowing them to be used in **ANALYSES** for the uses for which they are intended. Stable and consistent data.

Return on Investment (ROI)

A measure of the value that a project yields to its stakeholders.

Risk Assessment

The process of identifying (1) all assets an organization possesses, (2) all potential threats to those assets, (3) all points of vulnerability to those threats, (4) the probability of potential threats being realized, and (5) the cost estimates of potential losses. Risk assessment enables an organization to at least consider the range of potential threats and vulnerabilities it faces, and is the first step in effectively securing an information and technology system.

S

Sanitization of the Media

A process applied to data or storage media to make data retrieval unlikely for a given level of effort. Clear, Purge, and Destroy are actions that can be taken to sanitize data and media.

School Official

Any employee, including a teacher, who the school or district has determined to have a "legitimate educational interest" in the **PERSONALLY IDENTIFIABLE INFORMATION** from an **EDUCATION RECORD** of a **STUDENT**. School officials may also include third party contractors, consultants, volunteers, service providers, or another **PARTY** with whom the school or district has outsourced institutional services or functions for which the school or district would otherwise use employees under the school official exception in **FERPA**.

Score Distribution

How test scores are plotted along a scale.

Secure File Transfer Protocol

A broad term referring to network technology used to encrypt **AUTHENTICATION** information and data files in transit so that data files can be safely accessed, transferred, and managed.

Security Token

A physical (hardware or software) device that a **USER** possesses that serves to prove that the user requesting **ACCESS** is in fact who he or she claims to be. Examples of tokens include smart cards, key fobs, and USB keys.

SQL Injection

An attack technique exploiting security vulnerability of a website to gain access to its operations (e.g., to steal, delete, or modify the content of a database). It is often accomplished by inserting malicious SQL statements into user-input fields on a web application form.

Stakeholder Engagement

The process by which an organization or collection of organizations systematically involves its **STAKEHOLDERS** in its work.

Stakeholders

Individuals and groups directly or indirectly affected by a data system, beginning with the design, development, and implementation of the system. The groups include a wide range of individuals—from local education agencies (including administrators, teachers, parents, and **STUDENTS**), to the state and federal levels (including legislators, legislative staff, governors, etc.), to the public arena and business community.

Statewide Longitudinal Data System (SLDS)

A unit-level data system designed for collection, management, analysis, and reporting of statewide education data over time and across programs.

Student

Defined by **FERPA** regulations as any individual who is or has been in **ATTENDANCE** at an **EDUCATIONAL AGENCY OR INSTITUTION** and regarding whom the agency or institution maintains **EDUCATION RECORDS**.

Summative

Describes **ASSESSMENTS** used to evaluate learning, skill acquisition, or achievement at the conclusion of a defined instructional period.

Suppression

A process that involves removing data (e.g., from a cell or a row in a table) to prevent the **IDENTIFICATION** of individuals in small groups or those with unique characteristics. This method may often result in very little data being produced for small populations, and it usually requires additional suppression of non-sensitive data to ensure adequate protection of **PERSONALLY IDENTIFIABLE INFORMATION**.

Sustainability

The capacity to support a system or program over time with sufficient financial and human resources to meet current and future needs.

System Acceptance Testing

A phase in the software development lifecycle designed to ensure that the functionality and quality of the system are acceptable to end **USERS**.

System Deployment

Delivery, installation, and testing of a computer or system to put it in a state of operational readiness.

System Design

The means by which the operational needs of the data contributors and data **USERS** are translated into a technical infrastructure.

System Development

An iterative logical process that aims to create computer-coded or programmed software to address a unique business or personal objective, goal, or process. Software development is generally a planned initiative that consists of various steps or stages that result in the creation of operational software.

System Initiation/Planning

A phase in the data system lifecycle that includes documentation of the system's goals and the process that will be used to create and implement a technical solution to meet those goals.

System Requirements Analysis

A phase in the data system lifecycle in which the detailed specifications for the system are examined to determine whether they meet the stated end **USER** needs and goals for the system.

T

Targeted Request

A request for data in which the person requesting the information is trying to get information about a specific **STUDENT**. For example, if there was a rumor published in the local paper that a public official was disciplined for cheating during his senior year in high school, a request to the high school for the disciplinary records of students who were caught cheating during the year the public official was a senior would be considered a targeted request.

Technology Platform

The computer operating environment on which a data system will run.

Time Series

Data that are collected at regular intervals over a period of time. Time series data are useful for understanding **TRENDS**, anomalies, and change over time.

Timely

Data that are collected at the most useful time(s) (i.e., a "fall count" of students occurs during the fall season) and are expected to be submitted in a reasonable window of time, so that **ANALYSIS** and reporting can be most meaningful.

Transactional Processing System

A data system in which there is continuous updating of information through entries, deletions, and updates of data. In transactional processing, the data activities are processed one at a time. This is in contrast to batch processing, in which a group of activities is processed at one time.

Transformation Rules

The rules for converting a set of data values from the data format of the source data system into the data format of a different destination data system.

Transparency or Transparent Process

The essential condition for a free and open exchange whereby the rules and reasons behind business processes are clear and understood by all **STAKEHOLDERS**. Transparency involves a lack of hidden agendas and conditions, accompanied by the availability of full information required for collaboration, cooperation, and collective decisionmaking. It is characterized by visibility or accessibility of information, especially

concerning business practices, as well as the assurance that data being reported are accurate and are coming from the official source.

Transport Layer Security (TLS)

A **CRYPTOGRAPHIC** network protocol that provides **AUTHENTICATION**, **CONFIDENTIALITY**, and data integrity between two communicating applications. TLS is used as a mechanism to protect sensitive data during electronic dissemination across the Internet.

Trend

The general direction in which something is developing or changing.

Triangulation

A qualitative data **ANALYSIS** approach where three or more data points are utilized to answer research questions or to make generalizations about a **TREND**, group, or program in order to increase the credibility and **VALIDITY** of use of the findings.

U

Unauthorized Disclosure

Occurs when **PERSONALLY IDENTIFIABLE INFORMATION** from a **STUDENT'S EDUCATION RECORD** is made available to a third party who does not have legal authority to access the information. Such an unauthorized **DISCLOSURE** can happen inadvertently, as occurs when information about an individual is unintentionally revealed through, for example, a security breach of the electronic system that is used to maintain and access the education records, or when a teacher or administrator accidentally leaves paper reports that include personally identifiable information in an unsecured location. See also **AUTHORIZED DISCLOSURE**, **DISCLOSURE**, and **IDENTITY DISCLOSURE**.

Unique Identifier

A numeric or alphanumeric string that is associated with a single entity (e.g., a child, service provider, teacher, or local program) within a given system.

User

A person who accesses, analyzes, reports, and/or uses data.

User Interface

The visual part of a computer **APPLICATION** or operating system through which a **USER** interacts with a computer or software. The interface determines how commands are given to the computer or program and how information is displayed on the screen.

V

Valid-for-Use Data

Data that are sound for use in **ANALYSES**, in that they represent that which they purport to represent. Note: Data themselves are not valid; rather, data can be determined to be valid *for use*.

Validity

The degree to which an **ASSESSMENT**, program, or instrument measures what it purports to measure.

Variable

Any characteristic, number, or quantity that can be measured or counted.

Vision Statement

An aspirational description of how the system will help support the mid- and long-term education goals of the state. It should be heavily informed by the needs of a broad range of **STAKEHOLDERS**.

W

Web Portal

A website that serves as a gateway or a main entry point ("cyber door") on the Internet to a specific field of interest or an industry. A portal provides at least four essential services: (1) search engine(s), (2) email, (3) links to other related sites, and (4) personalized content. It may also provide facilities such as chat, a members list, or free downloads.

Workflow

A term used to describe the tasks, procedural steps, organizations or people involved, required input and output information, and tools needed for each step in a business process.